DOCUMENT RESUME

ED 347 885 HE 025 648

AUTHOR Bender, Timothy A.; Standage, Tamra Holmes
TITLE Influences on Students' Uses of Classroom Examination Feedback.
PUB DATE Apr 92
NOTE 13p.; Paper presented at the Annual Meeting of the American Educational Research Association (San Francisco, CA, April 20-24 1992).
PUB TYPE Reports - Research/Technical (143) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Academic Achievement; *College Students; Error Correction; *Feedback; Higher Education; Individual Characteristics; *Performance; Sex Differences; Student Attitudes; *Student Motivation; Teacher Response; *Tests

ABSTRACT

This study explored the effects of a combination of student and feedback characteristics on the postfeedback performance of Southwest Missouri State University undergraduate students in a classroom-like situation. The subjects were 26 male and 43 female students enrolled in 3 large sections of Introductory Psychology with all sections taught by the same female instructor. The procedure was as follows: (1) all students answered two questionnaires on the first day of class; (2) exams were administered approximately once every 3 weeks; (3) students were given a chance at voluntary feedback before the second exam; (4) feedback was provided to subjects in groups of 10 to 15 on the day following the second exam; and (5) 48 hours later, students received a readministration of the second exam. Results found a consistent pattern emerging for gender differences with males more motivated by grades, more likely to view exam feedback as a form of teacher control and generally more extrinsically motivated than females. It was also found that feedback was more effective for the higher- than the lower-achieving students. When data were reanalyzed with achievement level as a covariate gender differences disappeared. In addition the students with high expectations and a learning-orientation were particularly good users of feedback. Included are 11 figures and 9 references. (JB)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

Influences on Students' Uses of Classroom Examination Feedback

Timothy A. Bender
Southwest Missouri State University
Tamra Holmes Standage
Southwest Missouri State University

ED 347885

Objectives

Recent models of students' processing of exam feedback (Bangert-Drowns, Kulik, Kulik, & Morgan, 1991; Kulhavy & Stock, 1989) have included both student and feedback characteristics as influences on the effectiveness of exam feedback. This study was designed to explore the effects of a combination of student and feedback characteristics on postfeedback performance in a classroom-like situation.

Perspectives

We chose to investigate student characteristics for which existing research suggested a relationship with exam feedback. These included students' motivational responses to feedback (Bender and Horn, 1990; Bender, 1991), strategies of feedback use (Bender and Horn, 1990; Bender, 1991), achievement level (Bender, 1991; Gagne, et al., 1987), confidence in their responses (Kulhavy and Stock, 1989), senses of autonomy and control (Deci and Ryan, 1985; 1987), and orientations towards learning vs. grades (Eison, Pollio, and Milton, 1986.) On the basis of these studies, we expected effective feedback use to be positively related to attending to feedback about errors and guesses, relating feedback to one's own effort, performing well prior to receiving feedback, expecting high performance, and maintaining a a generally intrinsic, autonomous, learning-oriented approach to academics. We also expected feedback use to be negatively related to a generally extrinsic, control-oriented view of academics.

Our manipulation of the feedback was based on the work of McColskey and Leary (1985.) In a laboratory study, McColskey and Leary found that lower-achieving students benefited from self-referenced feedback, while higher-achieving students reacted more positively to norm-referenced feedback. Both feedback type and achievement level were randomly assigned. The source of the self-referenced information was a bogus verbal skills test and the task was solving difficult anagrams. We attempted a similar manipulation, but tried to keep the paradigm as close to regular classroom exam conditions as possible. In our study, feedback type was determined by a matching procedure based on actual classroom exam performance. Thus, achievement level was determined by actual performance. On the basis of the McColskey and Leary (1985) findings, we expected the low-scoring, self-referenced students to benefit more than the low-scoring norm-referenced students.

Method

Subjects Subjects were 26 male and 43 female undergraduates enrolled in three large sections of Introductory Psychology at Southwest Missouri State University. All sections were taught by the same female instructor. Subjects were required to participate in research as part of the course activities, but were provided a choice of projects in which to participate.

Procedure On the first day of class all students enrolled in these three sections of Introductory Psychology (n=414) completed the Reactions to Feedback (RF) (Bender & Horn, 1990) and the Learning Orientation/Grade Orientation (LOGO-II) (Eison, et al., 1986) questionnaires in a counterbalanced order.

Classroom exams were administered approximately once every three weeks. Five days prior to their second regular exam (Exam Two), the students were told they would have an opportunity to receive feedback about their performance on the upcoming exam. A sign-up sheet was distributed to each section. Subjects who volunteered were reminded immediately prior to Exam Two to bring their texts and notes to the feedback

---

Presented at the 1992 annual meeting of the American Educational Research Association in San Francisco.

#E 025 648

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

X This document has been reproduced as received from the person or organization originating it

☐ Minor changes have been made to improve reproduction quality

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy

1

2

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
Timothy A. Bender

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

BEST COPY AVAILABLE

sessions. They were then administered Exam Two in groups of 135 to 140. The exam consisted of 50 multiple-choice items. There were two forms of the exam distributed in such a manner that no two students with the same form were seated next to each other. Students were allowed 50 minutes to complete the exam. The exams and answer sheets were collected upon completion. The answer sheets were scored by computer. Immediately following the scoring of the exam, we assigned subjects to either the norm- or self-referenced feedback condition through a matching procedure based on their scores on Exam Two.

Feedback was provided to subjects in groups of 10 to 15 on the day following the exam. After entering a small classroom, subjects were asked to complete a short questionnaire on which they predicted their scores on the exam. Upon completion, the exams, unscored answer sheets, answer keys, and feedback slips were handed to the subjects. For the norm-referenced condition, the feedback slip provided the subjects with their raw scores and class percentiles. For the self-referenced condition, the feedback slips provided the subjects with their raw scores and the difference between their scores on Exam Two versus Exam One. Subjects were told to use their texts and class notes to review their exam performance. They were allowed 15 minutes for this review. At the end of the review period, the exams, answer sheets, keys, and feedback slips were collected. Subjects completed another short questionnaire on which they indicated to what degree they met their performance expectations on the exam. Subjects then were asked to complete the General Causality Orientation Scale (GCOS) (Deci & Ryan, 1985) and were told their responses would be compared to their exam scores. Upon completion of the GCOS, subjects were reminded to return in 48 hours to complete another questionnaire. The instructor was not present for the feedback or subsequent testing sessions.

When subjects returned 48 hours later, they were re-administered Exam Two. After finishing the exam, subjects completed a final questionnaire on which they were asked to indicate how much they had studied the tested material since the first administration of Exam Two. This provided a means of controlling for study outside of the experiment. Subjects were also asked to recall whether their feedback was norm- or self-referenced. This measure served as a manipulation check.

Results and Discussion

Only 5 subjects reported any additional study. No differences in performance were found between students who studied some additional time and those who did not. All reported means are harmonic means.

Gender Differences

One-way ANOVAs were used to test for gender differences. Significant differences occurred for the Teacher Control scale of the RF, $F(1,54)=8.96$, $p < .005$, $eta^2 = .14$, with means and standard errors of 10.35 (.513) and 8.51 (.339) for males and females, respectively. Males more often claimed that feedback made them feel as if the teacher was in control of their learning (Figure One.) Similarly, males were more control-oriented on the GCOS, $F(1,66)=8.47$, $p < .005$, $eta^2 = .11$, with means and standard errors of 54.88 (1.216) and 50.38

Insert Figure One about here.

(.957) for males and females, respectively (Figure Two.) Males were also more grade-oriented $F(1,54)=4.80$, $p <$

Insert Figure Two about here.

.04, $eta^2 = .08$, with means and standard errors of 44.41 (1.651) and 40.07 (1.09) for males and females, respectively (Figure Three.)

Insert Figure Three about here.

Although no gender differences were found for scores on the classroom administration of Exam Two, males performed poorer on the posttest (Figure Four), $F(1,67)=7.67$, $p < .008$, $eta^2 = .10$, with means and

Insert Figure Four about here.

standard errors of 36 (1.379) and 40.84 (1.072) for males and females, respectively. Males also were more likely to commit a different error $F(1,67)=4.65$, $p < .04$, $eta^2 = .06$, and less likely to correct their errors $F(1,67)=4.94$, $p < .03$, $eta^2 = .07$ (Figure Five.) The means and standard errors of the probabilities of committing different errors

Insert Figure Five about here.

for males and females are .14 (.018) and .09 (.014) for males and females, respectively. For the probability of correcting their errors, the means and standard errors for males and females are .60 (.035) and .70 (.028), respectively.

A consistent pattern emerged for gender differences. Males were more motivated by grades, adopted a greater general control-orientation, and were more likely to view exam feedback as a form of teacher control. All of these results suggest that males were more extrinsically motivated than females. This control-related extrinsic motivation has been related to poor exam performance in other studies (Deci & Ryan, 1987.) Similarly, in our results, males performed more poorly following feedback, and were poorer users of the feedback than were females.

Achievement Level

A median-split of the scores on the classroom exam allowed an investigation of the effects of achievement level on the tested material. A 2 (norm- vs. self-referenced feedback) x 2 (high vs. low performance on the first administration of Exam Two) x 2 (prefeedback vs. postfeedback performance on Exam Two) repeated measures ANOVA was used to test the hypothesis that lower-achieving students would benefit more from self-referenced feedback than from norm-referenced feedback. The between-subjects variables were the type of feedback and achievement level. The repeated measure was performance on each of the two administrations of the second exam. Significant main effects were found for achievement level $F(1,48)=66.36$, $p < .00009$, $eta^2 = .39$, and exam administration $F(1,48)=77.67$, $p < .00009$, $eta^2 = .18$. Across exams, higher-achieving students scored better than lower-achieving students, with means and standard errors of 41.1 (.84) and 31.4 (.84), respectively. Across achievement level, scores following feedback were greater than those from the first administration, with means and standard errors of 39.6 (.54) and 32.9 (.54), respectively. The two-way interaction between achievement-level and administration was also significant $F(1,48)=4.81$, $p < .04$, $eta^2 = .01$. Figure Six illustrates this interaction. All pairwise comparisons were significant. Both the higher- and

Insert Figure Six about here.

lower-scoring students improved their mean performance from the first administration to the second. However, students who performed better prior to feedback continued to perform better following feedback.

However, no significant effects involving norm- vs. self-referenced feedback were found. A manipulation check on the type of feedback indicated that the manipulation was not effective. Of the subjects who received norm-referenced feedback, 7 thought the feedback was self-referenced, 4 thought it was norm-referenced and 21 thought it was both. The manipulation was more effective for the self-referenced condition. Twenty-three of the students in the self-referenced conditioned correctly identified the type of feedback they received, while 12 thought they received both forms. Two subjects did not respond. We suspect that characteristics of the normal classroom setting complicate the manipulation of norm- vs. self-referenced feedback. First, to the student,

feedback in a classroom may be more important than feedback in a lab task. Second, in the classroom, students are generally not provided with bogus feedback, as they were in the McColskey and Leary (1985) study. Third, overlap of the treatment conditions may occur in the classroom. Norm-referenced feedback may have a self-referenced component if students recall their performance on previous exams. Self-referenced feedback may have a norm-referenced component if students compare their scores. In our study, as in most traditional classroom settings, there was no control over either of these potential confounds.

Single factor ANCOVAs were used with gender as the covariate to further explore the effect of achievement-level. Significant differences were found for postfeedback exam performance $F(1,63)=26.21$, $p < .00009$, $eta^2 = .26$, the proportions of same errors $F(1,63)=4.55$, $p < .04$, $eta^2 = .07$, different errors $F(1,63)=6.97$, $p < .02$, $eta^2 = .09$, new errors $F(1,63)=17.99$, $p < .0001$, $eta^2 = .21$, and corrected errors $F(1,63)=11.29$, $p < .002$, $eta^2 = .14$. In general, higher-achieving students on the first administration of Exam Two performed better on the postfeedback administration, committed a lower proportion of same, different, and new errors, and a greater proportion of corrected errors than did lower-achieving students (Figure Seven.) For the probability of

Insert Figure Seven about here.

committing same errors, the means and standard errors for the higher- and lower-achieving students were .19 (.025) and .27 (.026), respectively. For different errors, the results for higher- and lower-achieving students were .08 (.016) and .14 (.016), respectively. For new errors, the data for higher- and lower-achieving students were .60 (.053) and .93 (.055), respectively. Finally, the probabilities of correcting an error for higher- and lower-achieving students were .73 (.028) and .59 (.029), respectively.

Feedback was more effective for the higher- than the lower-achieving students. Higher-achieving students were more likely to correct their errors, and less likely to commit the same, different, or new errors. Initial exam performance is a powerful predictor of postfeedback performance. In our study differences in initial achievement accounted for 26% of the variance in postfeedback scores and 7% to 21% of the variance in different error patterns following feedback. Differences in initial performance accounted for more variance in postfeedback performance than did gender differences. In fact, when we re-analyzed the data to test for gender differences, using achievement level as a covariate, all gender differences disappeared. The strength of differences in initial achievement combined with its mediating effect prompts us to suggest that achievement should must be included in any complete model of feedback processing.

Other Variables

In order to assess the effects of students' expectations on their subsequent performance, we used a Chi^2 to compare students' predicted grade with whether they scored above or below the median (Figure Eight.)

Insert Figure Eight about here.

Students were relatively accurate in their predictions $Chi^2(4)=16.54$, $p < .003$. Higher-achieving students predicted mostly A's, B's, and C's. Lower-achieving students predicted mostly C's and D's.

In order to explore the relationship between students' performance expectations and their post-feedback performance, we formed three expectation levels. The first level was comprised of students who predicted A's and B's; the second, C's; and the third, D's and F's. Data were analyzed using ANCOVAs with gender and achievement level as covariates and predicted performance as the single independent variable. The only significant result occurred for new errors $F(2,60)=3.37$, $p < .05$, $eta^2 = .07$. Means and standard errors for the high-, middle-, and low-expectation groups were .63 (.06), .86 (.059), and .83 (.08), respectively. Fisher's LSD revealed that the only significant pair-wise comparison was between the high- and middle-expectations groups. However, as seen in Figure Nine, the low- and middle-expectation groups performed similarly. Students who

Insert Figure Nine about here.

expected the highest scores were less likely to commit new errors than were students who predicted lower performance, particularly those who predicted C exam scores. New errors occur when students fail to use feedback to confirm initially correct responses. Thus, independent of initial achievement level and gender, students with high expectations maintained their correct responses better than students with more moderate expectations. These results indirectly support Kulhavy and Stock's (1989) assertion that confidence affects students' processing of feedback.

The probability of committing new errors was also related to performance on the RF $F(1,39)=4.59$, $p < .04$, $eta^2 = .06$, and the LOGO-II $F(1,45)=4.38$, $p < .05$, $eta^2 = .06$, independent of achievement level and gender. However, no effects were found for the general causality orientation. As seen in Figure Ten, students who

Insert Figure Ten about here.

expressed a greater general use of feedback were less likely to commit new errors. Means and standard errors of the probabilities of new errors for students who scored high and low in the general use of feedback were .64 (.068) and .83 (.047), respectively. As seen in Figure Eleven, students who were more learning-oriented also

Insert Figure Eleven about here.

were less likely to commit new errors than students who scored lower in learning-orientation. Means and standard errors for students who scored high and low in learning-orientation were .66 (0.67) and .84 (0.51), respectively.

In summary, the data partially support our predictions. A general orientation to using feedback, prior achievement on the tested material, performance expectations, and a learning-oriented perspective on academics were positively related to feedback use. The strongest factor affecting students' use of feedback was their performance prior to the feedback. Higher-scoring students were better feedback processors, as reflected in their error patterns. Gender differences also appeared in which females were better feedback processors than the more extrinsic, control-oriented males. Once differences in achievement level were controlled, gender differences disappeared. Independent of both achievement level and gender, students who claimed to use feedback, those with high expectations, and those with a learning-orientation were particularly good users of feedback about questions they had answered correctly. However, we did not find improved performance when low performance was presented in self-referenced terms. Differences between the classroom and lab may partially account for the difference in results.

Implications The result that students who are higher-achieving prior to feedback maintain this advantage after feedback is not particularly surprising. However, the reaction of educators to this difference in performance is important. McColskey and Leary (1985) suggested that the correct response is to adapt the feedback to the student. We found that this suggestion was not easily accomplished in the classroom. An alternative response would be to attempt to influence how students react to the feedback. This might be accomplished through direct training in the use of feedback or indirectly by influencing motivational variables which affect the use of feedback. From a practical standpoint, it is likely that both the development of varied and more informative feedback procedures, as well as training students in the use of the feedback are central to better learning from exam feedback.

References

Bangert-Drowns, R. L., Kulik, C. C., Kulik, J. A., and Morgan M. (1991). The instructional effect of feedback in test-like events. Review of Educational Research, 61, 213-238.

Bender, T. A. (April, 1991). Achievement, response certitude, motivation, and feedback on classroom examinations. Presented at the annual meeting of the American Educational Research Association in Chicago.

Bender, T. A. and Hom, H. L. Jr. (June, 1990). Differences in motivation mediate processing of exam feedback. Presented as a poster session at the annual meeting of the American Psychological Society in Dallas.

Deci, E. L. and Ryan, R. M. (1985). The general causality orientations scale: Self-determination in personality. Journal of Research in Personality, 19, 109-134.

Deci, E. L. and Ryan, R. M. (1987). The support of autonomy and the control of behavior. Journal of Personality and Social Psychology, 53, 1024-1037.

Eison, J. A., Pollio, H. R., and Milton, O. (1986). Educational and personal characteristics of four different types of learning- and grade-oriented students. Contemporary Educational Psychology, 11, 54-67.

Gagne, E. D., Crutcher, R. J., Anzelc, J., Geisman, C., Hoffman, V. D., Schutz, P., and Lizcano, L. (1987). The role of student processing of feedback in classroom achievement. Cognition and Instruction, 4, 167-186.

Kulhavy, R. W. and Stock, W. A. (1989). Feedback in written instruction: The place of response certitude. Educational Psychology Review, 1, 279-308.

McColskey, W. and Leary, M. R. (1985). Differential effects of norm-referenced and self-referenced feedback on performance expectancies, attributions, and motivation. Contemporary Educational Psychology, 10, 275-284.

Figure 1. Gender Differences Teacher Control

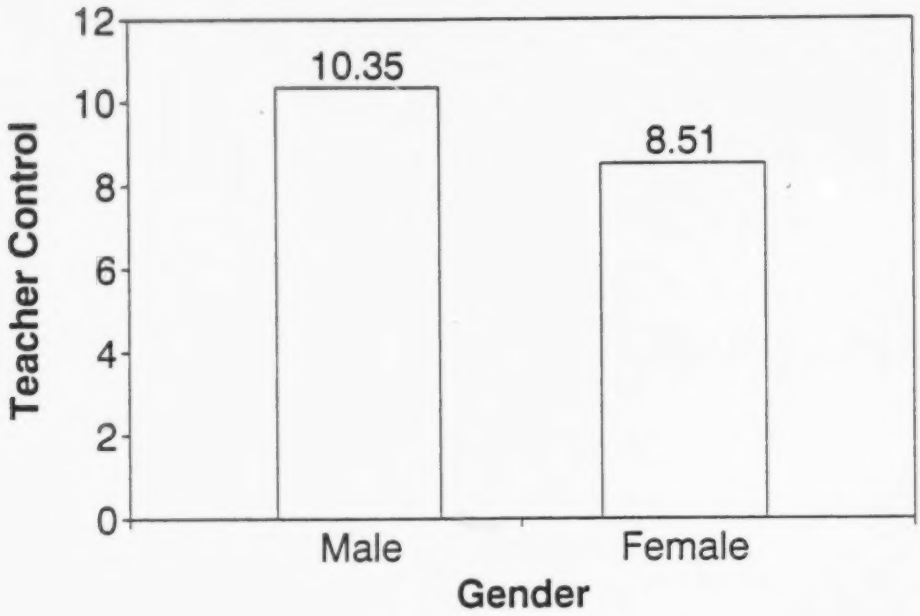

Figure 2. Gender Differences: Control Orientation

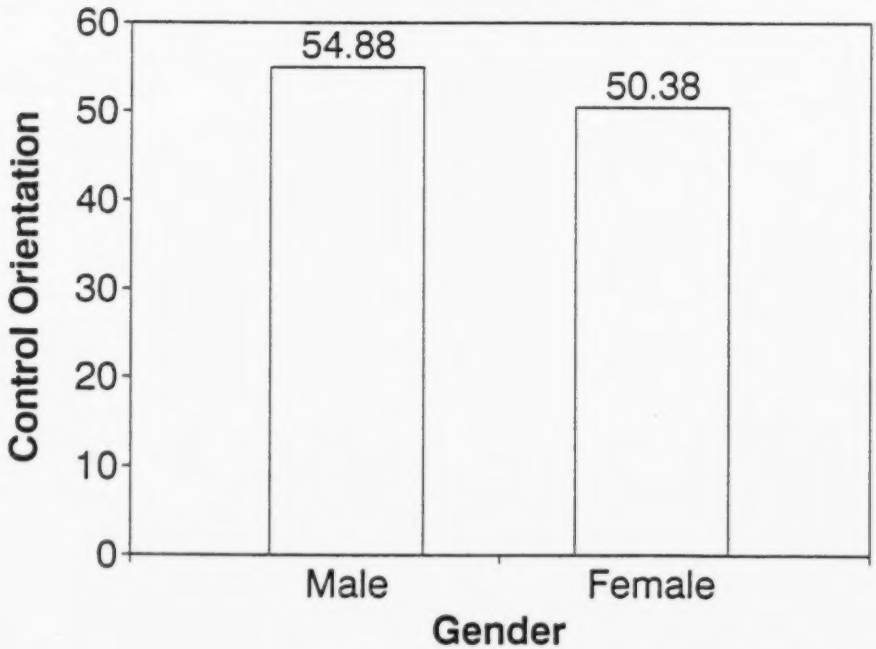

Figure 3. Gender Differences: Grade Orientation

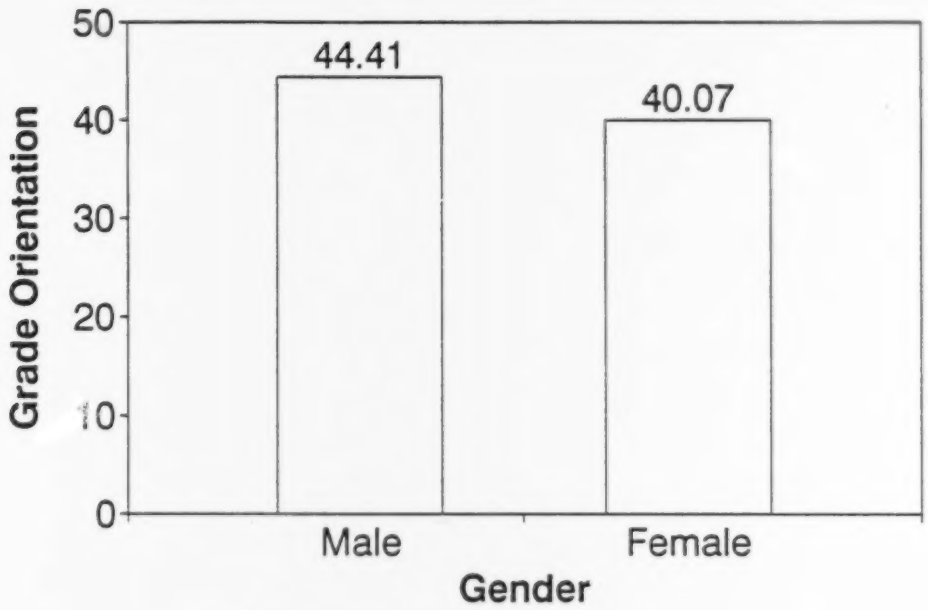

Figure 4. Gender Differences: Posttest Scores

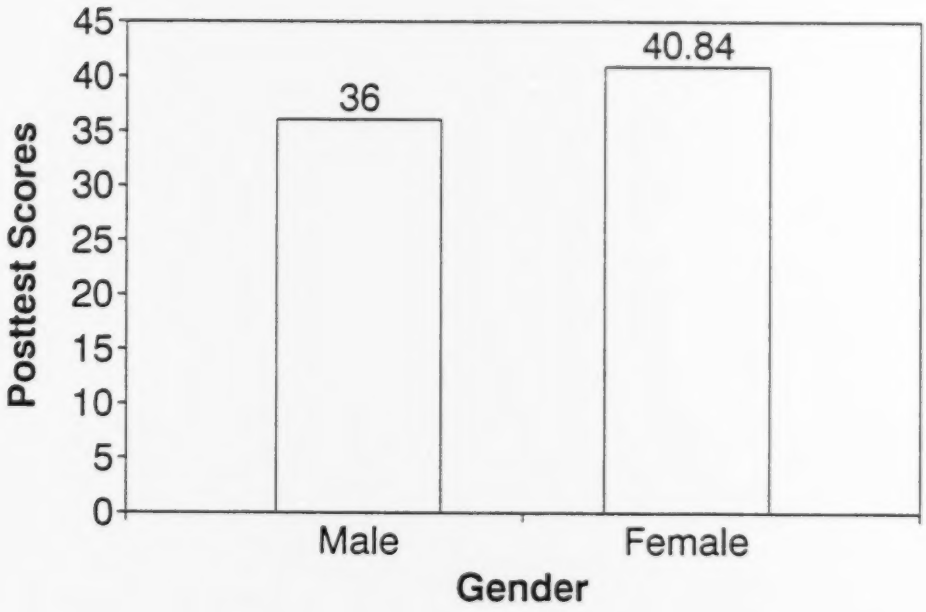

Figure 5. Gender Differences: Different and Corrected Errors

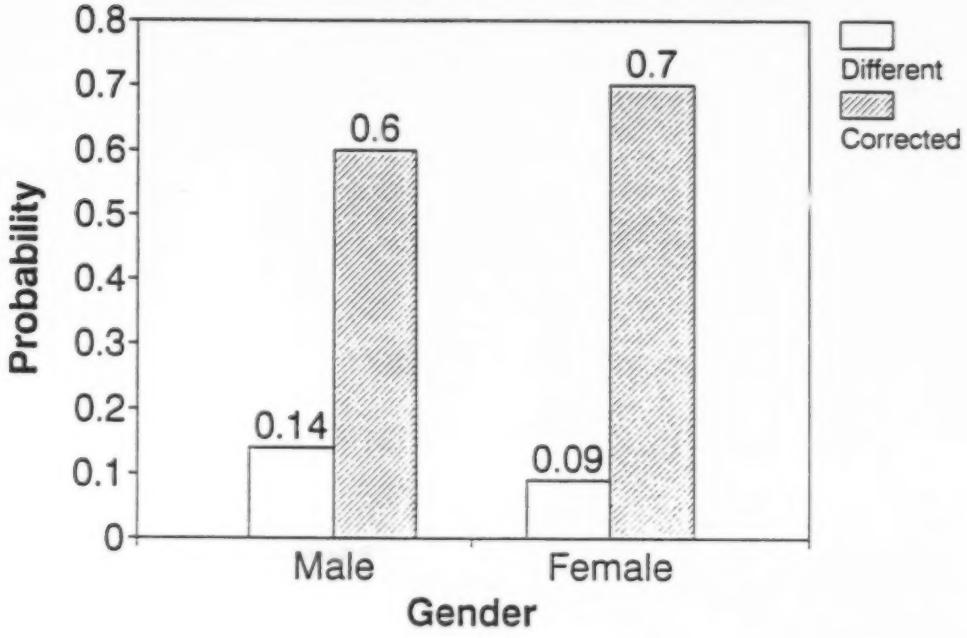

Figure 6. Achievement X Pre- and Posttest

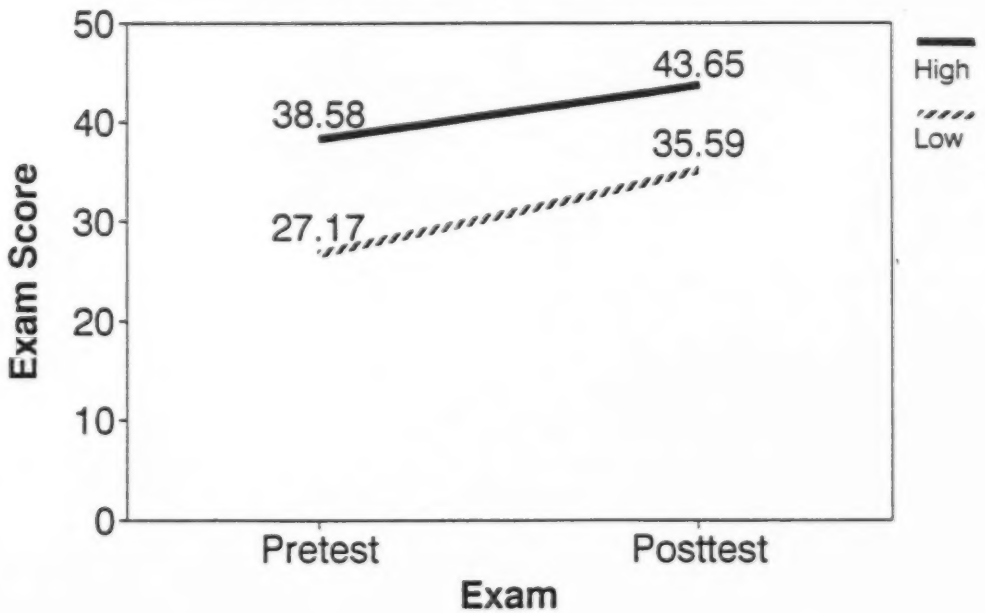

Figure 7. Error Patterns

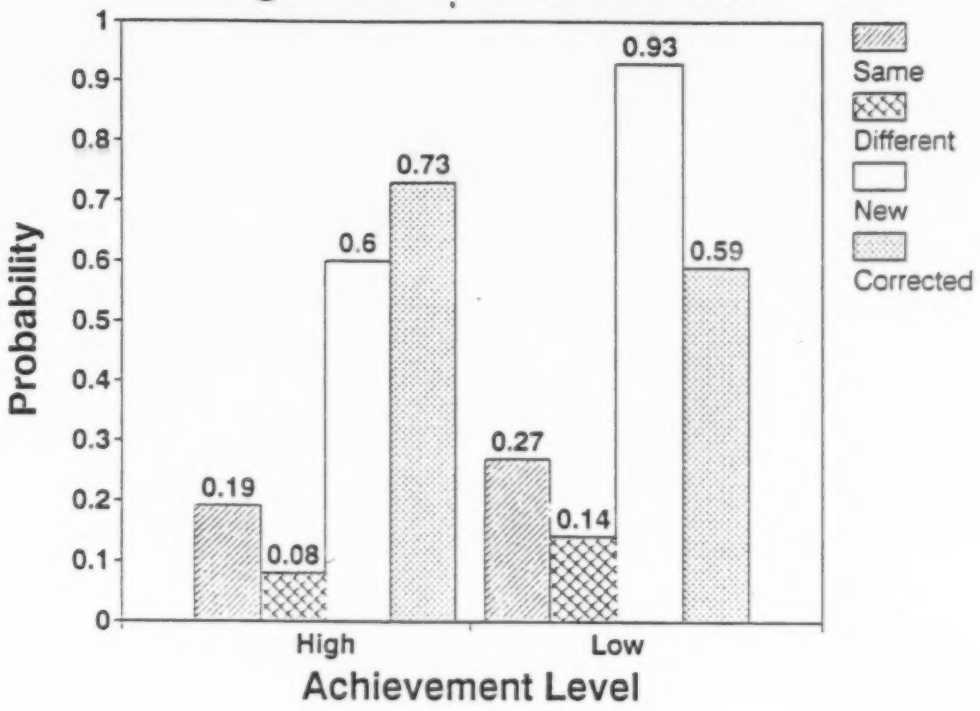

Figure 8. Predicted Grades

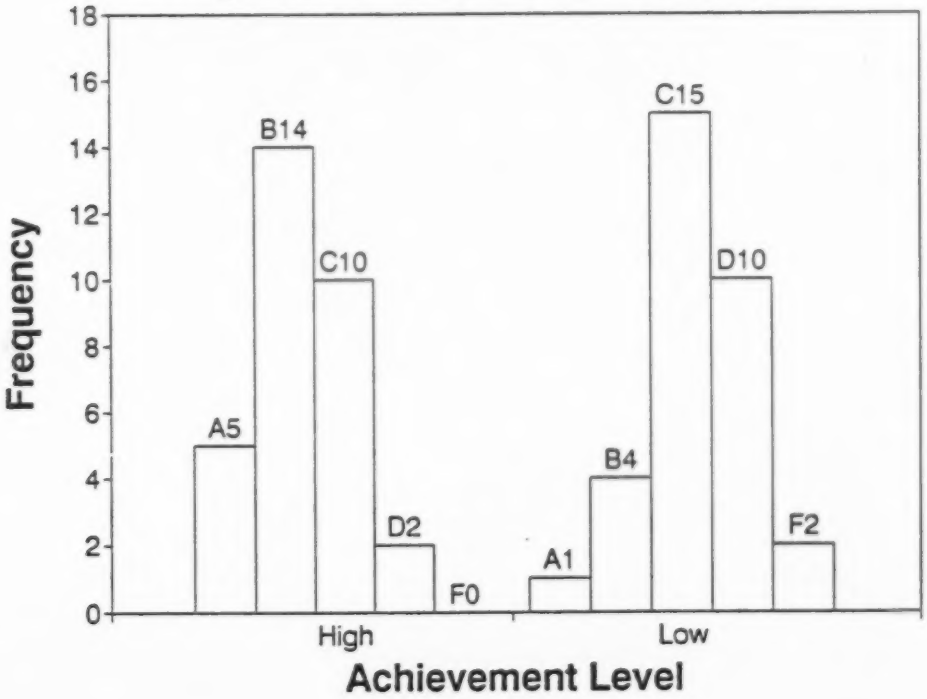

Figure 9. New Errors and Predicted Grades

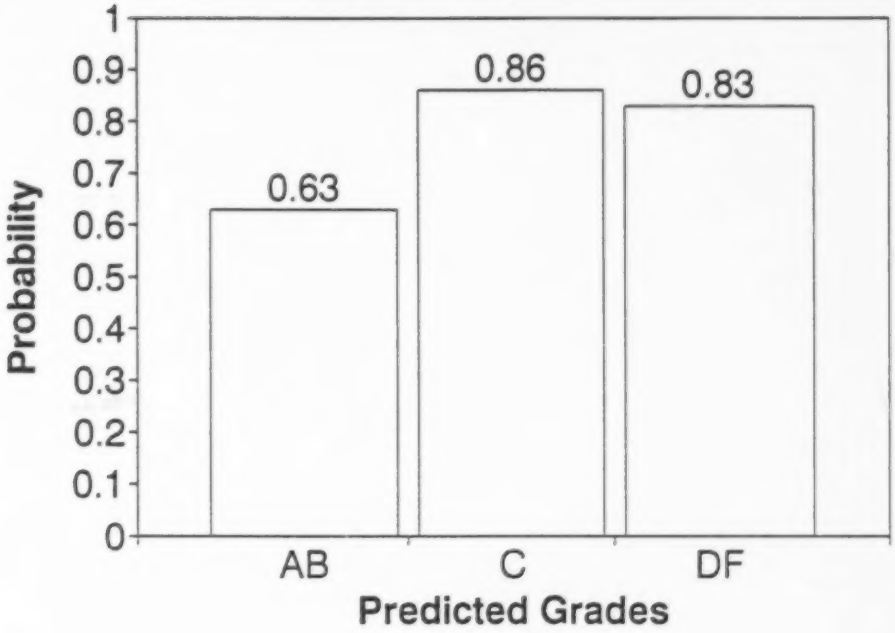

Figure 10. New Errors and General Use of Feedback

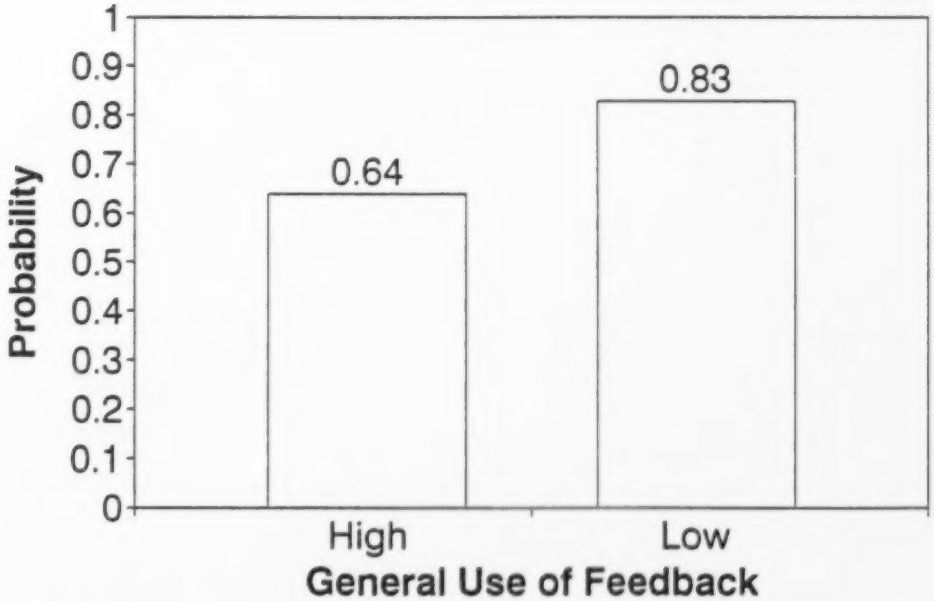

Figure 11. New Errors and Learning Orientation